



# Glycan Shield and Fusion Activation of a Deltacoronavirus Spike Glycoprotein Fine-Tuned for Enteric Infections

Xiaoli Xiong,<sup>a</sup> M. Alejandra Tortorici,<sup>b,c</sup> Joost Snijder,<sup>a</sup> Craig Yoshioka,<sup>d</sup> Alexandra C. Walls,<sup>a</sup> Wentao Li,<sup>e</sup> Andrew T. McGuire,<sup>f</sup> Félix A. Rey,<sup>b,c</sup> Berend-Jan Bosch,<sup>e</sup> David Veessler<sup>a</sup>

<sup>a</sup>Department of Biochemistry, University of Washington, Seattle, Washington, USA

<sup>b</sup>Institut Pasteur, Unité de Virologie Structurale, Paris, France

<sup>c</sup>CNRS UMR 3569 Virologie, Paris, France

<sup>d</sup>Department of Biomedical Engineering, Oregon Health and Science University, Portland, Oregon, USA

<sup>e</sup>Virology Division, Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

<sup>f</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

**ABSTRACT** Coronaviruses recently emerged as major human pathogens causing outbreaks of severe acute respiratory syndrome and Middle East respiratory syndrome. They utilize the spike (S) glycoprotein anchored in the viral envelope to mediate host attachment and fusion of the viral and cellular membranes to initiate infection. The S protein is a major determinant of the zoonotic potential of coronaviruses and is also the main target of the host humoral immune response. We report here the 3.5-Å-resolution cryo-electron microscopy structure of the S glycoprotein trimer from the pathogenic porcine deltacoronavirus (PDCoV), which belongs to the recently identified *Deltacoronavirus* genus. Structural and glycoproteomics data indicate that the glycans of PDCoV S are topologically conserved compared with the human respiratory coronavirus NL63 S, resulting in similar surface areas being shielded from neutralizing antibodies and implying that both viruses are under comparable immune pressure in their respective hosts. The structure further reveals a shortened S<sub>2</sub>' activation loop, containing a reduced number of basic amino acids, which participates in rendering the spike largely protease resistant. This property distinguishes PDCoV S from recently characterized betacoronavirus S proteins and suggests that the S protein of enterotropic PDCoV has evolved to tolerate the protease-rich environment of the small intestine and to fine-tune its fusion activation to avoid premature triggering and reduction of infectivity.

**IMPORTANCE** Coronaviruses use transmembrane S glycoprotein trimers to promote host attachment and fusion of the viral and cellular membranes. We determined a near-atomic-resolution cryo-electron microscopy structure of the S ectodomain trimer from the pathogenic PDCoV, which is responsible for diarrhea in piglets and has had devastating consequences for the swine industry worldwide. Structural and glycoproteomics data reveal that PDCoV S is decorated with 78 N-linked glycans obstructing the protein surface to limit accessibility to neutralizing antibodies in a way reminiscent of what has recently been described for a human respiratory coronavirus. PDCoV S is largely protease resistant, which distinguishes it from most other characterized coronavirus S glycoproteins and suggests that enteric coronaviruses have evolved to fine-tune fusion activation in the protease-rich environment of the small intestine of infected hosts.

**KEYWORDS** coronaviruses, cryo-EM, fusion proteins

Coronaviruses are large enveloped viruses, with single-stranded positive-sense RNA genomes, classified in four genera (*Alpha*-, *Beta*-, *Gamma*-, and *Deltacoronavirus*) based on their sequence similarity. Most recognized coronaviruses are animal viruses,

**Received** 14 September 2017 **Accepted** 23 October 2017

**Accepted manuscript posted online** 1 November 2017

**Citation** Xiong X, Tortorici MA, Snijder J, Yoshioka C, Walls AC, Li W, McGuire AT, Rey FA, Bosch B-J, Veessler D. 2018. Glycan shield and fusion activation of a deltacoronavirus spike glycoprotein fine-tuned for enteric infections. *J Virol* 92:e01628-17. <https://doi.org/10.1128/JVI.01628-17>.

**Editor** Tom Gallagher, Loyola University Medical Center

**Copyright** © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to David Veessler, [dveessler@uw.edu](mailto:dveessler@uw.edu).

For a companion article on this topic, see <https://doi.org/10.1128/JVI.01556-17>.

but four coronaviruses, namely, human coronavirus 229E (HCoV-229E), HCoV-OC43, HCoV-NL63, and HCoV-HKU1, are known to continuously circulate in the human population and are associated with up to 30% of respiratory tract infections (1). In addition, severe acute respiratory syndrome (SARS-CoV) and Middle East respiratory syndrome (MERS-CoV) coronaviruses are zoonotic viruses causing deadly pneumonia in humans (2). SARS-CoV and MERS-CoV have resulted in more than 8,000 and 2,000 cases, with fatality rates of 10 and 35%, respectively. No specific antiviral treatments or vaccines are approved for human coronaviruses, and zoonosis remains a great pandemic threat.

The abilities to recognize the appropriate receptor and to efficiently enter host cells are key requirements for cross-species spillover of zoonotic viruses such as influenza (3). For coronaviruses, these two functions are carried out by the spike (S) glycoprotein. Therefore, structural and functional studies of S glycoproteins can provide invaluable information to evaluate the cross-species transmission potential of these viruses. The coronavirus S protein is a class I viral fusion protein that forms homotrimers decorating the viral envelope. It is composed of an N-terminal  $S_1$  subunit, responsible for receptor binding, and a C-terminal  $S_2$  subunit, which contains the fusion machinery. The combined activities of the two subunits promote coronavirus attachment to host cells and subsequent fusion of the viral and cellular membranes via irreversible conformational changes, initiating viral infection. Since it is the major surface protein, S is also the main target of neutralizing antibodies during infection and a focus of vaccine design.

The zoonotic potential of coronaviruses is determined by the receptor-binding properties of the S protein. For instance, SARS-CoV and MERS-CoV bind with high affinity to their cognate human receptors, angiotensin-converting enzyme 2 (ACE2) and dipeptidyl peptidase 4 (DPP4), respectively (4, 5). Metagenomic data revealed that many MERS-CoV-like and SARS-CoV-like viruses exist in bats, and one such virus, WIV-1, isolated from bat feces, shares 99.9% nucleotide sequence identity with SARS-CoV. The S protein encoded by WIV-1 binds human, bat, and civet ACE2 orthologues, allowing the virus to efficiently infect human cells expressing any of these three orthologues (6, 7). Similarly, HKU4-CoV and HKU5-CoV, which are closely related to MERS-CoV, have been identified in bats, and HKU4-CoV can be modified to efficiently bind human DPP4 by substituting three amino acids in the S receptor-binding domain (8, 9).

The zoonotic potential of coronaviruses is further determined by fusion activation, which requires S processing by host proteases. Up to two cleavage sites are present in S glycoproteins: a site found at the boundary between the  $S_1$  and  $S_2$  subunits of some coronavirus S (the  $S_1/S_2$  site) and a conserved site upstream from the fusion peptide (the  $S_2'$  site) (10).

For a subset of coronaviruses, such as mouse hepatitis coronavirus (MHV), SARS-CoV, and MERS-CoV, the S glycoprotein is cleaved at the  $S_1/S_2$  junction during biogenesis and viral egress (10–13). This proteolytic event, along with subsequent binding to the host receptor, enhances processing at the  $S_2'$  site and participates in MERS-CoV or SARS-CoV fusion activation (11, 13). Moreover, substitution of two residues at the boundary between the  $S_1$  and  $S_2$  subunits enables efficient processing by human proteases and allows the bat-infecting HKU4-CoV S protein to mediate entry into human cells (14).

Proteolysis at the conserved  $S_2'$  site is essential for fusion activation of all characterized coronavirus S proteins, and it can occur at the host membrane or in internal cellular compartments. For instance, transmembrane protease/serine protease (TMPRSS) processing of SARS-CoV and MERS-CoV S at the cell membrane, furin-mediated processing of HCoV-NL63 and MERS-CoV S in the early endosomes, or endolysosomal protease-mediated triggering of SARS-CoV S (by cathepsin L) and MHV S are key events orchestrating spatial and temporal activation of fusion to ensure successful viral entry into host cells (12, 13, 15). Alternatively, porcine epidemic diarrhea coronavirus (PEDV), which replicates in the epithelial cells of the small intestine, undergoes S proteolytic activation by trypsin, which is highly abundant in the lumen of this organ (16). These examples illustrate how the

availability of host proteases and the mechanism of proteolytic activation can directly restrict coronavirus activation, viral tropism, and pathogenesis.

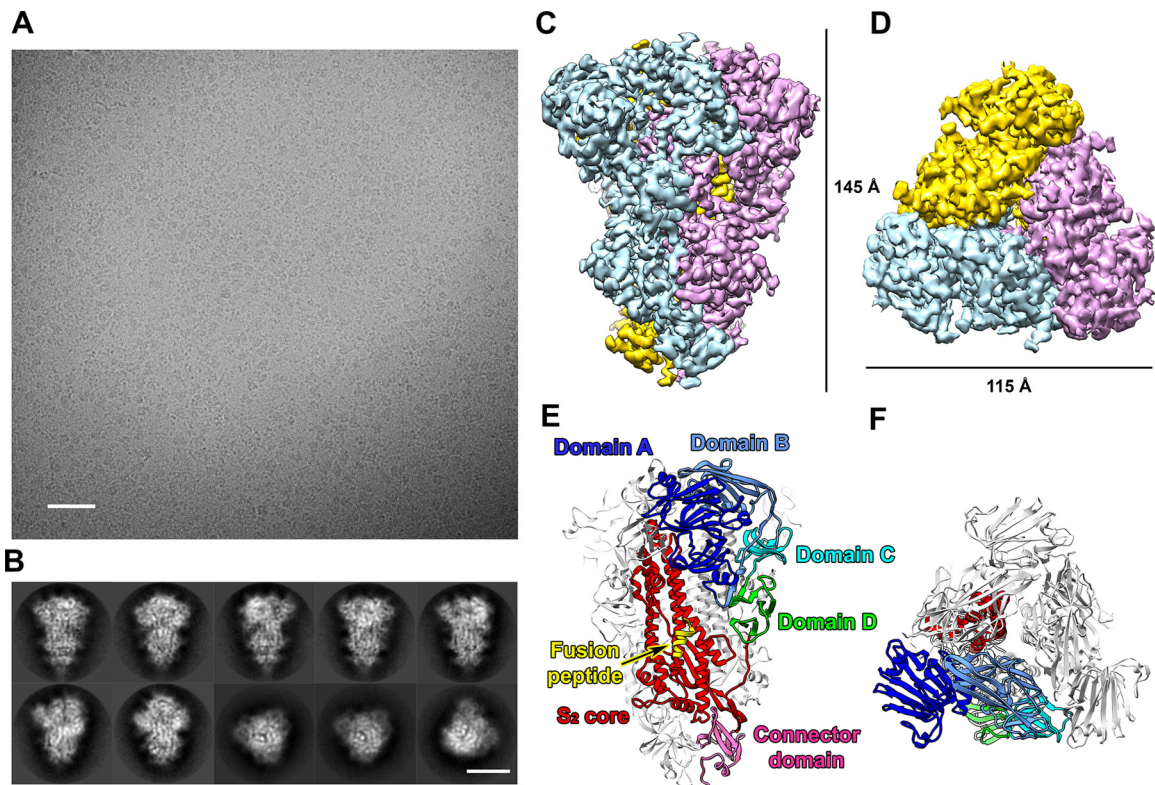
One common pattern shared by both SARS and MERS outbreaks is that although they both originated in bats, an intermediate host with closer physical proximity to humans allowed for more efficient cross-species transmission. Palm civets and camels were the most probable intermediate hosts for SARS-CoV and MERS-CoV, respectively (7, 17, 18). Due to their proximity to humans, pigs also acted as intermediate hosts for the influenza pandemic (19) and for the emergence of Nipah virus in Malaysia (20). To date, only alphacoronaviruses (hereinafter,  $\alpha$ -coronaviruses) and  $\beta$ -coronaviruses have been implicated in human diseases, and several S glycoproteins from viruses belonging to these two genera have been structurally characterized (21–26). To the best of our knowledge, no porcine coronaviruses have crossed the species barrier to infect humans, and their receptor usage appears to favor porcine orthologues. PEDV, however, can infect pig, human, monkey, and bat cells, suggesting it has the potential to spill over to species other than pig (27). As a result, cross-species transmission of coronaviruses poses an imminent and long-term threat to human health, which emphasizes the need for surveying and studying these viruses to prevent and control infections.

The recently emerged porcine deltacoronavirus (PDCoV) is responsible for diarrhea in piglets and has had devastating consequences for the swine industry worldwide (28, 29). No vaccines or treatments are available for PDCoV. Here, we report the cryo-electron microscopy (cryo-EM) structure of the PDCoV S trimer, revealing that it has a molecular architecture most closely related to that of the S glycoproteins of the *Alphacoronavirus* genus ( $\alpha$  genus). Integrating structural and glycoproteomics data, we discovered that PDCoV S masks potential epitopes with glycans in a way reminiscent of the human respiratory  $\alpha$ -coronavirus HCoV-NL63 S glycoprotein (22). These results support a relatedness between  $\alpha$ - and  $\delta$ -coronavirus S glycoproteins and suggest that the immune system of infected hosts exert comparable selection pressure on these viruses, which has led to these adaptations. The structure also reveals the C-terminal  $S_2$  fusion machinery of the PDCoV S protein features a short  $S_2'$  activation loop which appears to be largely resistant to proteolysis by trypsin/chymotrypsin. We conclude that PDCoV has evolved to be highly adapted to the protease-rich environment of the enteric tract to ensure proper spatial and temporal activation of fusion and prevent premature triggering, which would significantly impact virus infectivity.

## RESULTS

**Structure determination of the PDCoV S glycoprotein.** PDCoV was first identified in Hong Kong in 2012 (29), and it has since spread rapidly in the swine population across the globe (28, 29). Due to its recent emergence, relatively little is known about this virus compared to other swine coronaviruses. One feature that distinguishes PDCoV from other known coronaviruses is that it encodes one of the smallest S glycoproteins. We therefore set out to explore the architectural diversity of S proteins across coronavirus genera to understand shared and unique features of the structurally uncharacterized  $\delta$  genus.

We used *Drosophila melanogaster* S2 cells to produce the PDCoV/USA/Illinois121/2014 S ectodomain (residues 1 to 1098) with a C-terminal fusion adding a GCN4 trimerization motif and a Strep-tag (30). Following sample vitrification by triple blotting (31), data were acquired on an FEI Titan Krios electron microscope equipped with a Gatan Quantum GIF energy filter operated in zero-loss mode and a Gatan K2 Summit electron-counting camera operated in super-resolution mode (Fig. 1A and B). We determined a three-dimensional (3D) reconstruction at 3.5-Å resolution, resolving most amino acid side chains, disulfide bonds, and N-linked glycans (see Fig. S1A in the supplemental material). These features were used as fiducials to confirm the sequence register during model building (Fig. 1C to F and Fig. S1B to E). Starting from the HCoV-NL63 S structure (22), we obtained an atomic model of the PDCoV S trimer using manual modeling in Coot (32) and Rosetta density-guided iterative refinement (33). The final model comprises residues 52 to 1021 and 21 N-linked glycans (Table 1).



**FIG 1** Cryo-EM structure of the PDCoV S protein. (A) A representative micrograph of vitreous ice-embedded PDCoV S protein at 3.4- $\mu$ m defocus. Scale bar, 510 Å. (B) Selected 2D class averages of the PDCoV S protein. Scale bar, 85 Å. (C and D) Side (C) and top (D) views of the PDCoV S cryo-EM map filtered at 3.5-Å resolution and sharpened with a B-factor of  $-150 \text{ Å}^2$ . The density is colored for each protomer. (E and F) Ribbon representation of the PDCoV S trimer structure rendered with the same orientations as those in panels C and D. One protomer is colored according to the indicated structural domains, whereas the other two protomers are colored gray.

The PDCoV S protein assembles as a compact trimer with a height of  $\sim 145 \text{ Å}$  and a width of  $115 \text{ Å}$  (Fig. 1C and D). The  $S_1$  subunit has a modular organization comprising four distinct domains, designated A, B, C, and D, whereas the  $S_2$  subunit adopts a mostly helical elongated architecture with a connector domain appended to its C-terminal end (21, 22) (Fig. 1E and F).

**The extensive PDCoV S glycan shield.** The unsharpened PDCoV S map resolves 21 N-linked glycans for each protomer that form prominent protrusions extending from

**TABLE 1** Data collection and refinement statistics

Parameter	Value
Data collection	
No. of particles	263,800
Pixel size (Å)	0.665
Voltage (kV)	300
Electron dose ( $e^-/\text{Å}^2$ )	40
Refinement	
Resolution (Å)	3.5
Map-sharpening B factor (Å <sup>2</sup> )	$-150$
Model validation	
Favored rotamers (%)	98.7
Poor rotamers (%)	0.35
Ramachandran allowed (%)	99.79
Ramachandran outliers (%)	0.21
Clash score	1.41
MolProbity score	1.1



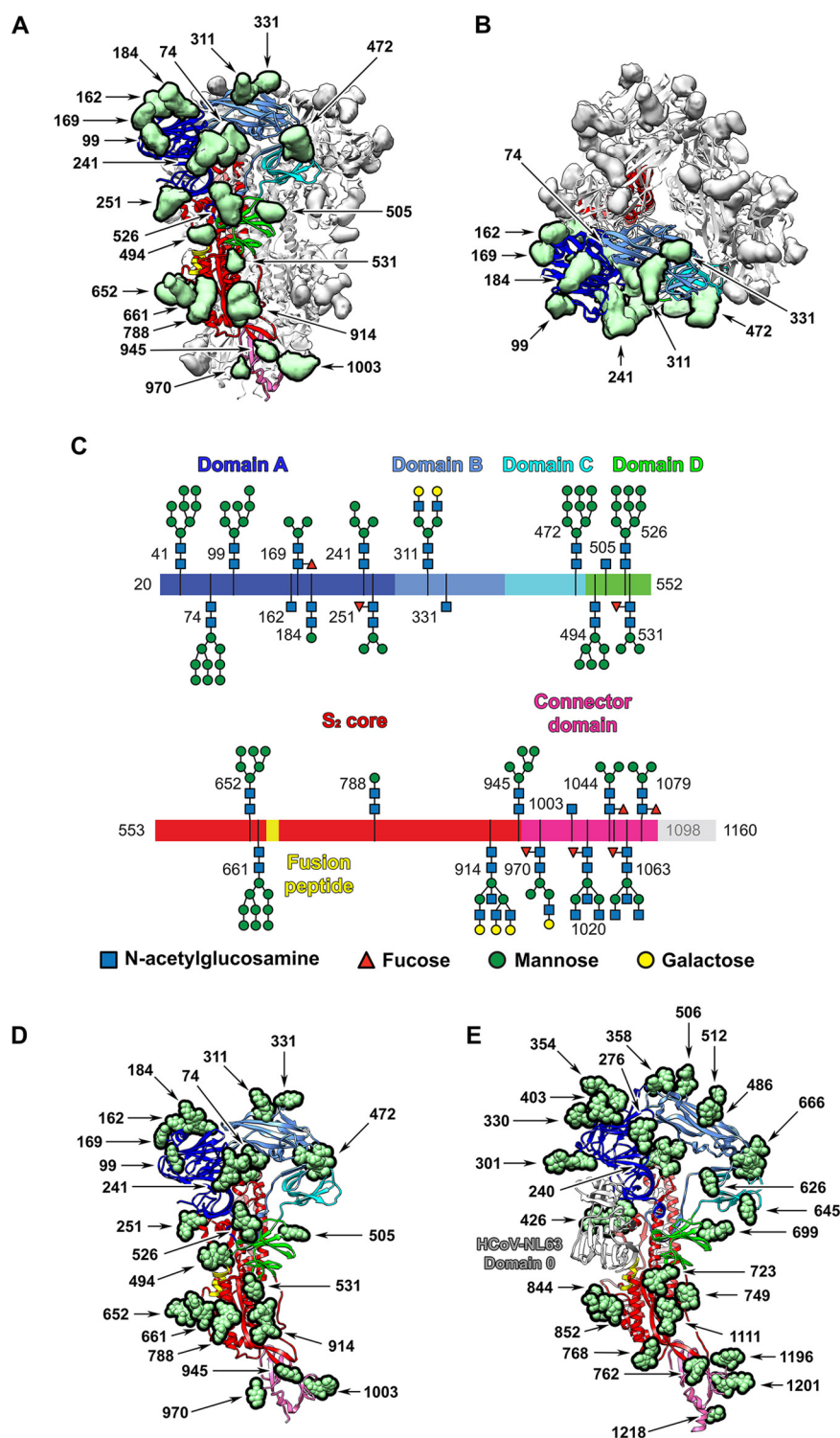
the protein surface (Fig. 2A and B and Fig. S1F and G). Using on-line reverse-phase liquid chromatography with electron transfer/high-energy collision-dissociation tandem mass spectrometry (LC-MS/MS) (34), we detected 16 N-linked glycosylation sites corresponding to those observed in the cryo-EM map and confirmed 5 additional sites located in the structurally unresolved N- and C-terminal parts of the protein (Fig. 2C and Table S1). Combining our structural and mass spectrometry data, we found evidence for glycosylation at 26 out of 27 possible NXS/T glycosylation sequons. The intact glycopeptides detected by MS/MS for PDCoV S expressed in *Drosophila* S2 cells corresponded mostly to paucimannosidic glycans containing 3 mannose residues (with or without core fucosylation) and oligomannose glycans containing 4 to 9 mannose residues. We also detected complex glycans (with or without core fucosylation), which appears compatible with the accessibility and crowding of these carbohydrate chains that would permit processing (35, 36).

Overall, the glycan coverage of PDCoV S is dense and extensively decorates the accessible surface of the trimer. Although we detected substantially more N-linked glycans for HCoV-NL63 S (34 sites per protomer) (22), 6 validated glycans reside within the N-terminal domain 0, which is absent from PDCoV S and explains most of the discrepancy in the number of sites. Strikingly, numerous glycans identified in the PDCoV S structure overlap glycans in the HCoV-NL63 S protein, either strictly or topologically, with most differences toward the viral membrane-distal end of the molecule (Fig. 2D and E). Transmission of zoonotic viruses into humans can result in drastic changes in glycosylation, as exemplified by the human influenza H3 hemagglutinin that has doubled its number of glycosylation sites since the 1968 pandemic, although its amino acid sequence remains ~88% identical (37). There is considerable sequence divergence between the HCoV-NL63 and PDCoV S glycoproteins, which share 43% amino acid sequence identity. The observation that numerous glycosylation sites are conserved between the two proteins suggests that  $\alpha$ - and  $\delta$ -coronaviruses face similar immune pressure in their respective hosts, and that the areas that are masked by the conserved glycans are key to the function of these S glycoproteins. Based on the information gained from the HCoV-NL63 S structure (22), for which glycans appear to contribute to masking the receptor-binding loops from antibody recognition, we suggest that the glycan shield of PDCoV S and other coronavirus S glycoproteins could assist in immune evasion similarly to the well-characterized HIV-1 envelope trimer (35, 36).

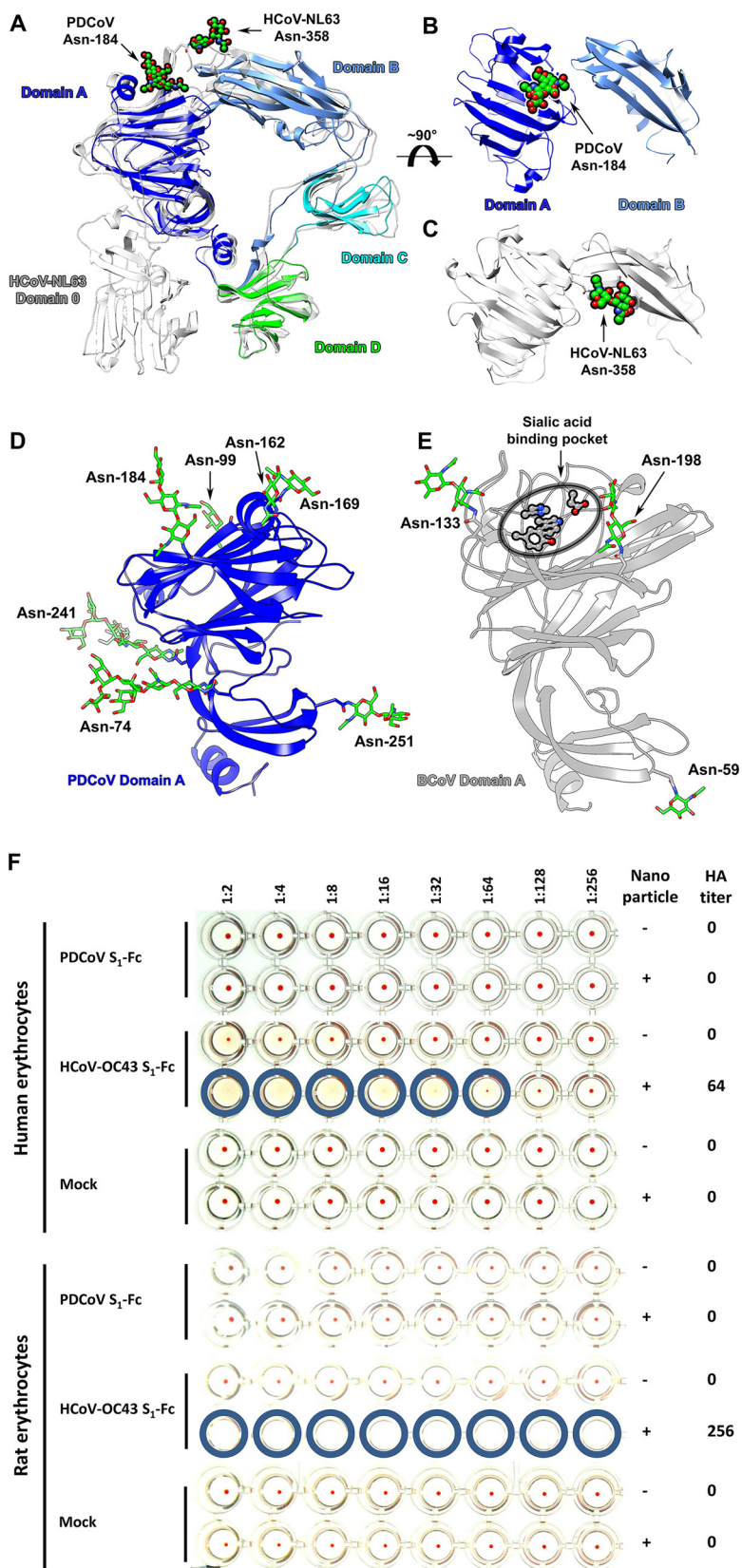
Finally, coronavirus S glycans have previously been proposed to participate in host cell entry (38), since L-SIGN lectin can be used as an alternative receptor by SARS-CoV (39) and HCoV-229E (40), and it is conceivable they play a similar role for other S proteins.

**Architecture of the S<sub>1</sub> receptor-binding subunit.** The PDCoV and HCoV-NL63 S<sub>1</sub> subunits exhibit strikingly similar structures (root mean square deviations [RMSD] of 2.7 Å over 448 aligned C $\alpha$  positions), except for the absence of the N-terminal domain 0 in the former glycoprotein (Fig. 3A). Deletion of domain 0, which is responsible for attachment to sialoglycans, from the porcine transmissible gastroenteritis virus (TGEV) S gene gave rise to porcine respiratory coronavirus (PRCV) and in turn resulted in a loss of enteric tropism (41, 42). PDCoV and HCoV-NL63, however, exhibit opposite behavior, as they target the enteric or the respiratory tract despite the absence or presence, respectively, of a domain 0 in their S glycoproteins. We describe below the functionally relevant similarities and differences detected in the PDCoV S structure relative to other coronavirus S structures.

Domain A is located at the viral membrane-distal side and accounts for a large part of the exposed surface area of the S<sub>1</sub> subunit. It folds as a galectin-like  $\beta$ -sandwich supplemented with a helix on the viral membrane-distal side and a three-stranded antiparallel  $\beta$ -sheet plus a helix on the proximal side. The domain A surface is heavily glycosylated and features 7 glycans for PDCoV (Fig. 3D). We previously reported that the HCoV-NL63 S glycan linked to Asn358 (domain A) points toward the receptor-



**FIG 2** Glycosylation profile of the PDCoV S protein. (A and B) Two orthogonal views of the PDCoV S trimer rendered as ribbons. Glycan density extracted from the unsharpened reconstruction is colored green for one protomer and gray for the other two protomers. Labels indicate the position of N-linked glycosylated asparagine residues. (C) Schematic summary of all detected N-linked glycans. Each site shows the most extensive glycan structure detected, either by mass spectrometry or cryo-EM. A full overview of all detected N-linked glycans is provided in Table S1 in the supplemental material. Glycan moieties are represented as symbols according to the key, and the structural domains are individually colored and indicated in a linear representation of the PDCoV S sequence. (D and E) Ribbon representation of PDCoV (D) and HCoV-NL63 (E) S protomers with glycans visualized by cryo-EM shown as green spheres.



**FIG 3** Structural features of the PDCoV S<sub>1</sub> subunit and the galectin-like domain A. (A) Superposition of the PDCoV and HCoV-NL63 S<sub>1</sub> subunits highlights the absence of domain 0 in PDCoV S<sub>1</sub>. (B) View of the interface between PDCoV S<sub>1</sub> A and B domains showing the Asn-184 glycan points away from domain B. (C) View of the interface between HCoV-NL63 S<sub>1</sub> A and B domains showing the Asn-358 glycan contributes

(Continued on next page)

binding domain B, masking residues involved in receptor recognition. A marked difference between the A domains of PDCoV S and HCoV-NL63 S is that the  $\beta$ -hairpin harboring Asn358 in HCoV-NL63 S features a deletion of 10 residues, significantly shortening it in PDCoV S (Fig. 3A to C). Moreover, the topologically equivalent glycan linked to residue Asn-184 of PDCoV S is protruding away from domain B and does not significantly cover it, in contrast to what was observed for HCoV-NL63 (Fig. 3A to C).

OC43, HKU1, and bovine coronavirus (BCoV) are known to use 9-O-acetyl-sialylated cellular receptors for attachment to host cells (43, 44). Structural and biochemical studies showed that domain A mediates these interactions and mapped key residues involved (25, 45), and a nanoparticle-displayed multimeric OC43 S<sub>1</sub> subunit exhibited high hemagglutination titer (Fig. 3F). Comparison of the PDCoV, HKU1, and BCoV domain A structures indicated PDCoV cannot interact with 9-O-acetyl-sialoglycans in a similar way due to the absence of the strictly conserved residues involved in binding (BCoV Tyr162, Glu182, Trp184, and His185) and of the loops delineating the binding cavity (Fig. 3D and E). In line with this observation, isolated or nanoparticle-displayed multimeric PDCoV S<sub>1</sub> subunit failed to interact with sialic acid using an erythrocyte hemagglutination assay (Fig. 3F), indicating that sialic acid (or at least the types of sialoglycans displayed on these erythrocytes) does not participate in PDCoV S attachment to host cells.

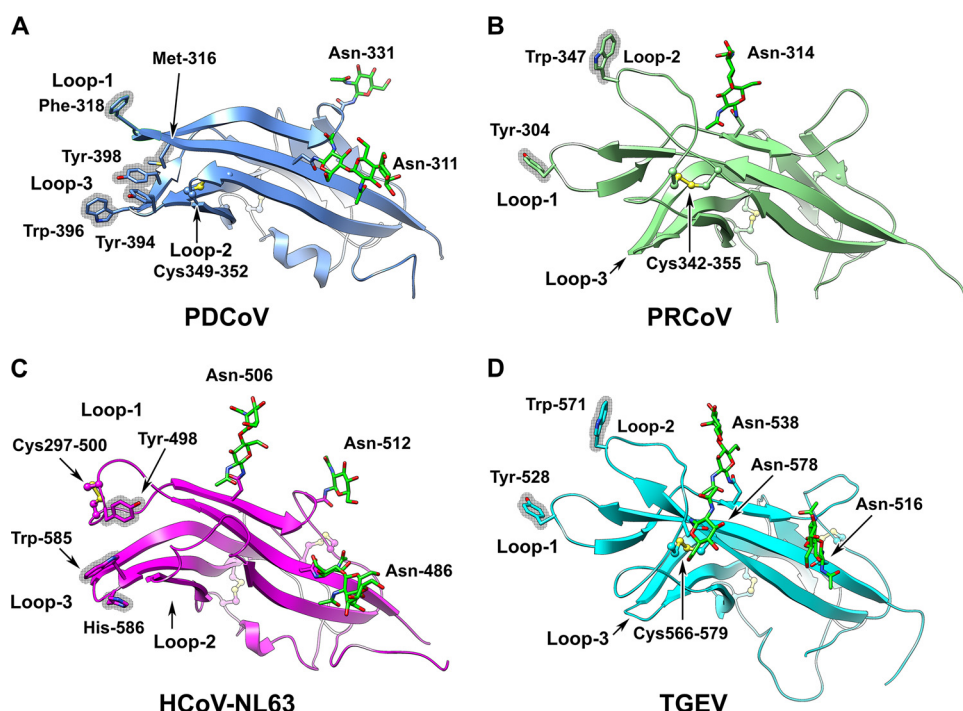
Domain B folds as a  $\beta$ -sandwich reminiscent of the equivalent domain of  $\alpha$ -coronaviruses, such as PRCV (RMSD, 2.1 Å over 108 aligned C $\alpha$  positions), HCoV-NL63 (RMSD, 1.9 Å over 107 aligned C $\alpha$  positions), and TGEV (RMSD, 3.0 Å over 109 aligned C $\alpha$  positions) (Fig. 4A to D) (22, 46, 47). The two PDCoV S glycosylation sites identified at Asn-311 and Asn-331 in domain B are topologically or strictly conserved with the HCoV-NL63 S glycans linked to Asn-486 and Asn-512, respectively. PRCV and TGEV B domains also feature topologically similar glycosylation sites on the solvent-exposed surface of the  $\beta$ -sandwich, and these glycans are likely to limit the immune response against this domain, which is known to be the target of neutralizing antibodies for several coronaviruses (47–52). The glycan linked to Asn-506 in HCoV-NL63 S is absent from PDCoV S, for which the equivalent residue is Ser-325 (Fig. 4A and C). Since masking of receptor-binding residues has been suggested to assist HCoV-NL63 immune evasion (22), the reduced overall glycan coverage of PDCoV domain B could result from weaker immune pressure directed at the receptor-binding region in pigs compared to HCoV-NL63 S in humans.

Previous work showed that the loops located at the viral membrane-distal end of the  $\beta$ -sandwich of domain B in  $\alpha$ -coronavirus S glycoproteins are responsible for binding to diverse host receptors, such as ACE2 (HCoV-NL63) (46) or pAPN (PRCV/TGEV) (47). Although the distal loops are significantly shorter for PDCoV than for these three  $\alpha$ -coronaviruses, loop 1 and loop 3 contain several aromatic residues (Fig. 4A to D). Since aromatic residues in these loops have been shown to directly participate in receptor binding for HCoV-NL63, PRCV, and TGEV, we speculate that they could also mediate interactions of the PDCoV B domain with its receptor. As is the case for HCoV-NL63 S, the PDCoV B domain has an opposite orientation, related by an  $\sim 180^\circ$  rotation, to the equivalent domain of  $\beta$ -coronavirus S glycoproteins (21). This results in burying the distal loops of the  $\beta$ -sandwich through interactions with domain A belonging to the same protomer and in turn restrain

### FIG 3 Legend (Continued)

to masking the receptor-binding loops. (D) Ribbon representation of PDCoV domain A. (E) Ribbon representation of BCoV domain A oriented identically to panel D. Highly conserved residues involved in sialic acid recognition are shown in ball-and-stick representation. Glycans are rendered as spheres in panels A to C or sticks in panels D and E and colored by atom type (carbon, green; nitrogen, blue; oxygen, red). (F) The PDCoV S<sub>1</sub> subunit C-terminally tagged with the Fc portion of human IgG (S<sub>1</sub>-Fc) was tested for its hemagglutination potential of an erythrocyte suspension of human or rat origin, either alone or premixed with protein A-coupled nanoparticles to increase the avidity of S<sub>1</sub>-Fc proteins for sialic acids. The sialic acid-binding S<sub>1</sub> subunit of HCoV-OC43 (GenBank accession no. [AA01015.1](https://www.ncbi.nlm.nih.gov/nuclseq/AA01015.1)) C-terminally fused to human Fc portion was used as a positive control. Mock indicates the condition where no S<sub>1</sub> subunit was used (negative control). Wells positive for hemagglutination are encircled.



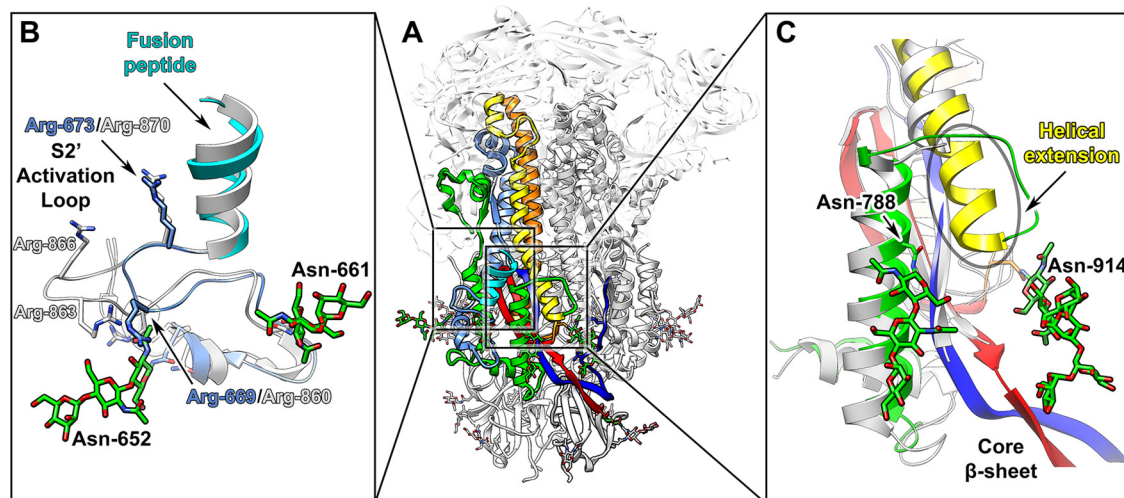


**FIG 4** Structural comparison of  $\alpha$ - and  $\delta$ -coronavirus receptor-binding domains. (A to D) Ribbon rendering of the putative receptor-binding domain B of the  $\delta$ -genus PDCoV S (A) and  $\alpha$ -genus PRCV S (B), HCoV-NL63 S (C), and TGEV S (D). Loops that have been implicated in receptor binding for  $\alpha$ -coronaviruses are indicated. Key aromatic residues that have been shown to take part in  $\alpha$ -coronavirus receptor binding and putatively involved in  $\delta$ -coronavirus receptor binding are highlighted. Disulfide bonds that stabilize receptor-binding loops are indicated, and glycans within the domain are shown as sticks (carbon, green; nitrogen, blue; oxygen, red).

the availability of the putative receptor-binding motif to interact with the receptor (Fig. 3B and C). As a result, it is likely that PDCoV and HCoV-NL63 S glycoproteins can undergo conformational changes similar to those described for domain B of SARS-CoV and MERS-CoV S to interact with their cognate receptors (23, 24, 26). A major difference, however, is that  $\beta$ -coronavirus S using domain B as a receptor-binding domain appears to spontaneously undergo these rearrangements, whereas  $\alpha$ - and  $\delta$ -coronavirus S do not and rely on an unidentified stimulus (22).

**Organization of the  $S_2$  fusion machinery.** The C-terminal  $S_2$  subunit trimer fuses the viral and cellular membranes at the onset of infection and is the most conserved region among coronavirus S glycoproteins. The PDCoV  $S_2$  subunit is structurally similar to  $\alpha$ - and  $\beta$ -coronavirus  $S_2$  subunits such as HCoV-NL63 (22) (RMSD, 1.7 Å over 413 aligned C $\alpha$  positions) and MHV (21) (RMSD, 2.2 Å over 291 aligned C $\alpha$  positions). The conserved architecture of the  $S_2$  fusion machinery across multiple genera highlights that coronaviruses rely on a common fusion mechanism to enter host cells (53). Despite these striking similarities, coronavirus fusion machineries exhibit differences with key functional implications for their activation mechanism and potential for zoonotic spillover.

The  $S_2'$  activation loop, which connects the upstream helix to the fusion peptide and regulates the spatial and temporal activation of fusion, is resolved in the PDCoV S cryo-EM map (Fig. 5A and B), as was the case for the HCoV-NL63 S (22) and SARS-CoV S structures (23, 24). However, the PDCoV  $S_2'$  loop is short (LTTRIGGR) and comprises 6 and 3 fewer residues than the HCoV-NL63 S (LPQRNIRSSRIAGR) and SARS-CoV S (ILPDPLKPTKR) counterparts, respectively. In addition, the  $S_2'$  loop of these viruses contains multiple positive charges, including two putative furin cleavage sites for HCoV-NL63 S, whereas the PDCoV  $S_2'$  loop harbors a single positively charged residue (Arg-669) in addition to the conserved Arg-673 residue (Fig. 5A and B). These structural features allow rationalizing the known protease requirements for fusion activation of the HCoV-NL63 S glycoprotein, which is preferentially cleaved by furin in the endo-

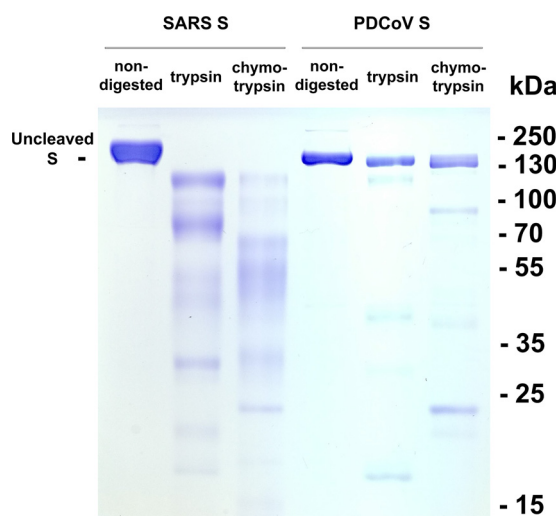


**FIG 5** Structural features of the PDCoV  $S_2$  subunit. (A) Ribbon representation of the PDCoV S trimer with the  $S_2$  subunit core of one protomer colored from blue to red (from N terminus to C terminus). (B) Zoomed-in view of the  $S_2'$  activation loop region. Two glycans, linked to Asn-669 and Asn-673, which are strictly conserved in HCoV-NL63 S, are shown as sticks (carbon, green; nitrogen, blue; oxygen, red). For comparison, the equivalent residues in the HCoV-NL63 S protein are indicated in gray. (C) The PDCoV S glycoprotein features an insertion of 14 amino acid residues in HR1 compared to the  $\beta$ -coronavirus MHV S protein, folding as an extended loop and a helical extension of two turns. The residues accounting for this HR1 insertion interact with the complementary insertion in HR2 in the postfusion conformation (see Fig. S2B).

somes, and of the SARS-CoV S glycoprotein, which is preferentially processed by cathepsin L in the endolysosomes, and explain the fact that trypsin-like TMPRSS proteases also can trigger both proteins (10, 15). The paucity of positive charges in the PDCoV  $S_2'$  trigger loop is in line with the requirement for trypsin or other pancreatic proteases to allow virus passaging and the fact that PDCoV is exposed to high concentrations of such proteases in the enteric tract of infected pigs (28).

Studies on influenza hemagglutinin highlighted that glycans can modulate cleavage site accessibility to proteases and in turn influence fusion activation (54, 55). Similar observations were drawn from comparisons between the MERS-CoV and HKU4 S glycoproteins (14). Notably, the PDCoV S glycans linked to Asn-652 and Asn-661 decorate the trimer surface near the  $S_2'$  trigger loop and could limit accessibility to proteases and play a role in orchestrating fusion activation (Fig. 5B). These glycans are conserved with an identical structural organization in HCoV-NL63 S and may have the same putative function (22).

Sequence alignment of representative S glycoproteins from viruses of the four coronavirus genera shows that  $\alpha$ - and  $\delta$ -coronaviruses feature a 14-residue-long insertion in the heptad repeat 1 (HR1) and HR2 motifs, corresponding to two heptad repeats, compared to  $\beta$ -coronaviruses (Fig. S2A).  $\gamma$ -Coronaviruses form a heterogeneous group comprising S glycoproteins without insertion but also S with one (BeCoV-SW1 and BdCoV-HKU22) or two (TurkeyCoV-MG10) additional heptad repeats in HR1 and in HR2 compared to  $\beta$ -coronaviruses. The HR1 insertion is resolved in the PDCoV S cryo-EM map (Fig. 5C, residues 797 to 811) and corresponds to the addition of two helical turns (also visible in the HCoV-NL63 S structure) preceded by a loop (poorly resolved in the HCoV-NL63 S structure). This polypeptide segment is known to refold to form a central triple-helical coiled coil in the postfusion S structure (53). The HR2 insertion cannot be visualized, as this region is disordered in the PDCoV S reconstruction and in all other coronavirus prefusion S structures. Mapping the HR1 and HR2 insertions on the HCoV-NL63 S postfusion core X-ray structure (56), however, reveals that these polypeptide segments directly interact within the 6-helix bundle (Fig. S2B). This suggests that the strict correlation of their presence or absence in both HR1 and HR2, along with the observation that insertions always correspond to an integer number of heptad repeats, is necessary to maintain the proper geometry of the fusion machinery and allow the conserved conformational changes driving membrane merger to take place with high efficiency.



**FIG 6** PDCoV S glycoprotein is resistant to digestive enzymes. Purified SARS S (1 mg/ml) and PDCoV S (0.5 mg/ml) glycoproteins were incubated with 0.1 mg/ml trypsin or chymotrypsin for 2 h at 22°C. The digestion reactions were analyzed on a 12% SDS-PAGE gel. After incubation, the SARS S protein was extensively proteolyzed, whereas a large fraction of the PDCoV S protein remains intact.

## DISCUSSION

Structural and functional studies of coronavirus S glycoproteins are key to understanding host and tissue tropism as well as the mechanisms of receptor binding and fusion activation. The data reported in the manuscript establish a strong connection between  $\alpha$ - and  $\delta$ -coronavirus S glycoproteins. HCoV-NL63, PRCV, TGEV, and PDCoV B domains fold as similar  $\beta$ -sandwiches that are structurally distinct from the single  $\beta$ -sheet observed for the equivalent domain of  $\beta$ -coronaviruses (4, 5, 21, 25, 47). Moreover, the structures of HCoV-NL63 S and PDCoV S show that both glycoproteins share an organization of their  $S_1$  subunits in which the B domain directly interacts with domain A from the same subunit to potentially limit accessibility of the receptor-binding loops to neutralizing antibodies. Sequence analysis indicates a strict correlation of the presence or absence of the HR1/HR2 insertions in the S glycoprotein sequence and an apparent evolutionary pressure restricting insertions/deletions to heptad repeat units, which we postulate are necessary for efficient S refolding and fusion. Based on these criteria, we put forward that  $\alpha$ - and  $\delta$ -coronavirus S glycoproteins share closer evolutionary relationships with each other than they do with S of the other two coronavirus genera, although insertions in HR1 and HR2 also have been detected in a subset of  $\gamma$ -coronavirus S proteins.

We previously recapitulated *in vitro* the proteolytic activation of MHV, SARS-CoV, and MERS-CoV prefusion S trimers via trypsin incubation under limited proteolysis conditions, which led to spontaneous refolding into postfusion  $S_2$  trimers (the ground state of the fusion reaction) (53). In contrast, the PDCoV S glycoprotein remained largely uncleaved even after extended incubation times with up to a 5:1 molar ratio of S to trypsin or chymotrypsin (0.1 mg/ml) (Fig. 6). These results suggest that fusion activation of PDCoV S, which is believed to be promoted by trypsin (28), involves an additional step to expose the  $S_2'$  cleavage site, such as the receptor-binding-induced conformational changes described for MERS-CoV (13), SARS-CoV (57), MHV (58), and PEDV (16). These distinct protease sensitivities are reminiscent of the differences reported between clinical isolates (CV777) and cell culture-adapted (caDR13) PEDV strains for which infectivity strictly requires or is hampered by trypsin, respectively (16). We put forward that the S glycoprotein sequence and, in turn, the structure of PDCoV and PEDV CV777 have evolved to be resistant to pancreatic proteases to which both viruses are exposed in the enteric tract during infection. Fine-tuning of fusion activation likely is achieved by restraining access to the  $S_2'$  cleavage site until receptor binding occurs at the host

cell surface. This event could promote conformational changes exposing the  $S_2'$  site to allow processing by trypsin or other proteases with exquisite spatial and temporal coordination. In contrast, SARS-CoV, MERS-CoV, and MHV are not expected to be exposed to pancreatic proteases during the virus life cycle, and their S glycoproteins presumably did not evolve with this selection pressure, explaining their sensitivity to trypsin and chymotrypsin. In agreement with what has been postulated for SARS-CoV (57) and PEDV caDR13 (16), trypsin sensitivity could result in premature cleavage/triggering of the prefusion S trimer and attenuation of infectivity and viral fitness.

## MATERIALS AND METHODS

**Plasmids.** A gene fragment encoding the PDCoV S ectodomain (residues 20 to 1098; UniProt accession code [W8Q9Y7](#)) was PCR amplified from a plasmid containing the full-length S gene. The PCR product was ligated to a gene fragment encoding a GCN4 trimerization motif (LIKRMKQIEDKIEIESKQK KIENEIARIKKIK) (21, 59, 60), an underlined thrombin cleavage site (LVPRGSL), an 8-residue-long Strep-tag (WSHPQFEK), and a stop codon, as previously described (30). Subsequent cloning was performed in the pMT<sub>B</sub>BiP<sub>V5</sub>His expression vector (Invitrogen) in frame with the *Drosophila* BiP secretion signal downstream of the metallothionein promoter.

A human codon-optimized gene encoding the ectodomain (residues 14 to 1180) of the SARS-CoV S protein (UniProt accession no. [P59594](#)) was cloned into a modified pOPING vector (61) (Addgene), introducing a C-terminal tobacco etch virus (TEV) protease cleavage site, a foldon, and a hexahistidine tag at the C terminus of the construct.

**Production of recombinant PDCoV S ectodomain in *Drosophila* S2 cells.** To generate a stable *Drosophila* S2 cell line expressing the recombinant PDCoV S ectodomain, we used Effectene (Qiagen) and 2  $\mu$ g of plasmid. Puromycin N-acetyltransferase was cotransfected as a dominant selectable marker. Stable PDCoV S-expressing cell lines were selected by addition of 7  $\mu$ g/ml puromycin (Invitrogen) to the culture medium 48 h after transfection. For large-scale production, the cells were cultured in spinner flasks and induced by 5  $\mu$ M CdCl<sub>2</sub> at a density of approximately 10<sup>7</sup> cells per ml. After a week at 28°C, clarified cell supernatants were concentrated 40-fold using Vivaflow tangential filtration cassettes (10-kDa cutoff; Sartorius) and adjusted to pH 8.0 before affinity purification using a Strep-Tactin superflow column (IBA), followed by gel filtration chromatography using a Superose 6 10/300 GL column (GE Life Sciences) equilibrated in 20 mM Tris-HCl, pH 7.5, and 100 mM NaCl. The concentration of the purified protein was estimated using absorption at 280 nm.

**Production of recombinant SARS S ectodomain in HEK293F cells.** Transient transfection of 250 ml HEK293F cells at a density of 10<sup>6</sup> cells/ml was performed using 293fectin (ThermoFisher) and Opti-MEM (ThermoFisher). After 3 days the cells were harvested before affinity purification with a Talon 5-ml cobalt column equilibrated in 25 mM sodium phosphate, pH 8.0, 300 mM NaCl, 10 mM imidazole. The purified protein was buffer exchanged into 20 mM Tris, pH 8.0, 150 mM NaCl and concentrated to 1.0 mg/ml.

**Cryo-EM specimen preparation and data collection.** Two microliters of purified PDCoV S at ~0.5 mg/ml was triple blotted (31) using 1.2/1.3 C-flat grids (Protochips), which had been glow discharged for 30 s at 20 mA. Grids were then plunge-frozen in liquid ethane using an FEI Mark I Vitrobot with 7.5-s blot time and an offset of -3 mm at 100% humidity and 25°C. Data were collected using SerialEM automatic data collection software (62) on an FEI Titan Krios operated at 300 kV and equipped with a Gatan GIF Quantum energy filter operated in zero-loss mode with a slit width of 20 eV and a Gatan K2 Summit direct electron detector camera operated in super-resolution mode. The dose rate was adjusted to ~5 counts/pixel/s, and each movie was acquired in counting mode fractionated in 70 frames of 200 ms. A total of 1,684 micrographs were collected in a single session using a defocus range between 1.5 and 4.0  $\mu$ m.

**Cryo-EM data processing.** Frame alignment was carried out using Motioncor2 (63). The parameters of the microscope contrast transfer function were initially estimated using GCTF (64) and then using CTFFIND4 (65). Particles were automatically picked using DoGPicker (66). Particle images were extracted and processed using Relion 2.0 (67) with a box size of 640 pixels<sup>2</sup> and a pixel size of 0.665 Å. Following reference-free 2D classification, we ran 3D classification with C1 symmetry (68) using an initial model generated with e2initialmodel.py in EMAN2. A total of 263,800 particles were selected to run a gold-standard 3D refinement imposing C3 symmetry using Relion 2.1 (69) that led to a map at 3.5-Å resolution. Postprocessing was done using Relion to apply an automatically generated B factor of -150 Å<sup>2</sup>. Reported resolutions are based on the gold-standard Fourier shell correlation (FSC) of 0.143 criterion (69, 70), and Fourier shell correlation curves were corrected for the effects of soft masking by high-resolution noise substitution (71). The soft mask used for FSC calculation had a 10-pixel cosine edge fall-off.

**Model building and analysis.** UCSF Chimera (72) was used to fit the HCoV-NL63 S structure (22) into the cryo-EM map before manual rebuilding in Coot (32, 73). Glycan density coming after an NXS/T motif was initially hand built into the density, where visible and glycan geometry then were refined using Rosetta, optimizing the fit to density as well as the energetics of protein/glycan contacts. The final model was refined using the symmetric modeling framework in Rosetta (33, 74). The quality of the final model was analyzed with Molprobity (75) and Privateer (76). All figures were generated with UCSF Chimera (72).

**Mass spectrometry.** A volume of 250 pmol of PDCoV S was incubated in a freshly prepared solution containing 100 mM Tris, pH 8.5, 2% sodium deoxycholate, 10 mM tris(2-carboxyethyl)phosphine, and 40



mM iodoacetamide at 95°C for 5 min, followed by 25°C for 30 min in the dark. Eighty pmol of denatured, reduced, and alkylated PDCoV S then was diluted into freshly made 50 mM ammonium bicarbonate and incubated for 14 h at 37°C either with 1:75 (wt/wt) trypsin (Sigma-Aldrich) or chymotrypsin (Sigma-Aldrich) or with alpha-lytic protease (Sigma-Aldrich). Formic acid then was added to a final concentration of 2% to precipitate the sodium deoxycholate in the samples, followed by centrifugation at 14,000 rpm for 20 min. The supernatant containing the (glyco)peptides was collected and spun again at 14,000 rpm for 5 min immediately before sample analysis. For each sample, 8  $\mu$ l was injected on a Thermo Scientific Orbitrap Fusion Tribrid mass spectrometer. A 35-cm analytical column and a 3-cm trap column filled with ReproSil-Pur C18AQ 5  $\mu$ m (Macherey-Nagel) beads were used. Nanospray LC-MS/MS was used to separate peptides over a 110-min gradient from 5% to 30% acetonitrile with 0.1% formic acid. A positive spray voltage of 2,100 was used with an ion transfer tube temperature of 350°C. An electron transfer/higher-energy collision-dissociation ion fragmentation scheme (34) was used with calibrated charge-dependent ETD parameters and supplemental higher-energy collision dissociation energy of 0.15. A resolution setting of 120,000 with an automatic gain control (AGC) target of  $2 \times 10^5$  was used for MS1, and a resolution setting of 30,000 with an AGC target of  $1 \times 10^5$  was used for MS2. Data were searched with Protein Metrics Byonic software (77) using a small custom database of recombinant protein sequences, including several coronavirus spike proteins, other viral glycoproteins, and the proteases used to prepare the glycopeptides. Reverse decoy sequences were also included in the search. Specificity of the search was set to C-terminal cleavage at R/K (trypsin), F/W/Y/M/L (chymotrypsin), or T/A/S/V (alpha lytic protease), allowing up to two missed cleavages, with EThcD fragmentation (b/y- and c/z-type ions). We used a precursor mass and product mass tolerance of 12 ppm and 24 ppm, respectively. Carbamidomethylation of cysteines was set as a fixed modification and methionine oxidation as a variable modification, and all four software-provided N-linked glycan databases were combined into a single nonredundant list used to identify glycopeptides. All glycopeptide hits were manually inspected, and only those with quality peptide sequence information are reported here.

**Proteolysis of PDCoV S and SARS S glycoproteins.** Proteins at a concentration of 0.5 mg/ml (PDCoV S) or 1 mg/ml (SARS-CoV S) were incubated with 0.1 mg/ml of either trypsin (Sigma Aldrich) or chymotrypsin at 22°C for 2 h. This reaction then was used for analysis by SDS-PAGE.

**Accession number(s).** The mass spectrometry data have been deposited in the PRIDE database with accession code [PXD007107](https://www.ebi.ac.uk/pride/archive/simpleSearch?q=PXD007107&submit=Search) (<https://www.ebi.ac.uk/pride/archive/simpleSearch?q=PXD007107&submit=Search>) and includes the raw data, Byonic search results, and the databases used for protein sequences and N-linked glycan modifications. The EM map and Protein Data Bank (PDB) model have been deposited under accession codes EMD-7094 (<http://www.emdatabank.org/>) and 6BFU, respectively.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JVI.01628-17>.

**SUPPLEMENTAL FILE 1**, PDF file, 6.0 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.1 MB.

## ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of General Medical Sciences under award numbers 1R01GM120553 (D.V.) and T32GM008268 (A.C.W.), the Seattle Structural Genomics Center for Infectious Disease from the National Institute of Allergy and Infectious Diseases under contract number HHSN272201700059C (D.V.), and a Pew Biomedical Scholars Award (D.V.). M.A.T. and F.A.R. acknowledge support from CNRS and Institut Pasteur recurrent funding. J.S. acknowledges support from the Netherlands Organization for Scientific Research (NWO; Rubicon 019.2015.2.310.006) and the European Molecular Biology Organization (EMBO; ALTF 933-2015). This work was also partly supported by the University of Washington's Proteomics Resource (UWPR95794).

We thank John Skehel for his critical reading of the article.

Microscopy was performed at the Multi-scale Microscopy Core (MMC) at Oregon Health & Science University (OHSU) with technical support from the OHSU-FEI Living Lab and the OHSU Center for Spatial Systems Biomedicine (OCSSB).

## REFERENCES

1. Zumla A, Chan JF, Azhar EI, Hui DS, Yuen KY. 2016. Coronaviruses—drug discovery and therapeutic options. *Nat Rev Drug Discov* 15:327–347. <https://doi.org/10.1038/nrd.2015.37>.
2. Vijay R, Perlman S. 2016. Middle East respiratory syndrome and severe acute respiratory syndrome. *Curr Opin Virol* 16:70–76. <https://doi.org/10.1016/j.coviro.2016.01.011>.
3. Schrauwen EJ, Fouchier RA. 2014. Host adaptation and transmission of influenza A viruses in mammals. *Emerg Microbes Infect* 3:e9. <https://doi.org/10.1038/emi.2014.9>.
4. Li F, Li W, Farzan M, Harrison SC. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309:1864–1868. <https://doi.org/10.1126/science.1116480>.

5. Lu G, Hu Y, Wang Q, Qi J, Gao F, Li Y, Zhang Y, Zhang W, Yuan Y, Bao J, Zhang B, Shi Y, Yan J, Gao GF. 2013. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. *Nature* 500:227–231. <https://doi.org/10.1038/nature12328>.
6. Menachery VD, Yount BL, Jr, Sims AC, Debbink K, Agnihothram SS, Gralinski LE, Graham RL, Scobey T, Plante JA, Royal SR, Swanstrom J, Sheahan TP, Pickles RJ, Corti D, Randell SH, Lanzavecchia A, Marasco WA, Baric RS. 2016. SARS-like WIV1-CoV poised for human emergence. *Proc Natl Acad Sci U S A* 113:3048–3053. <https://doi.org/10.1073/pnas.1517719113>.
7. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, Zhang YJ, Luo CM, Tan B, Wang N, Zhu Y, Cramer G, Zhang SY, Wang LF, Daszak P, Shi ZL. 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503:535–538. <https://doi.org/10.1038/nature12711>.
8. Yang Y, Du L, Liu C, Wang L, Ma C, Tang J, Baric RS, Jiang S, Li F. 2014. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. *Proc Natl Acad Sci U S A* 111:12516–12521. <https://doi.org/10.1073/pnas.1405889111>.
9. Wang Q, Qi J, Yuan Y, Xuan Y, Han P, Wan Y, Ji W, Li Y, Wu Y, Wang J, Iwamoto A, Woo PC, Yuen KY, Yan J, Lu G, Gao GF. 2014. Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. *Cell Host Microbe* 16:328–337. <https://doi.org/10.1016/j.chom.2014.08.009>.
10. Millet JK, Whittaker GR. 2015. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res* 202:120–134. <https://doi.org/10.1016/j.virusres.2014.11.021>.
11. Belouzard S, Chu VC, Whittaker GR. 2009. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A* 106:5871–5876. <https://doi.org/10.1073/pnas.0809524106>.
12. Burkard C, Verheije MH, Wicht O, van Kasteren SI, van Kuppeveld FJ, Haagmans BL, Pelkmans L, Rottier PJ, Bosch BJ, de Haan CA. 2014. Coronavirus cell entry occurs through the endo-/lysosomal pathway in a proteolysis-dependent manner. *PLoS Pathog* 10:e1004502. <https://doi.org/10.1371/journal.ppat.1004502>.
13. Millet JK, Whittaker GR. 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc Natl Acad Sci U S A* 111:15214–15219. <https://doi.org/10.1073/pnas.1407087111>.
14. Yang Y, Liu C, Du L, Jiang S, Shi Z, Baric RS, Li F. 2015. Two mutations were critical for bat-to-human transmission of Middle East respiratory syndrome coronavirus. *J Virol* 89:9119–9123. <https://doi.org/10.1128/JVI.01279-15>.
15. Bosch BJ, Bartelink W, Rottier PJ. 2008. Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide. *J Virol* 82:8887–8890. <https://doi.org/10.1128/JVI.00415-08>.
16. Wicht O, Li W, Willems L, Meuleman TJ, Wubolts RW, van Kuppeveld FJ, Rottier PJ, Bosch BJ. 2014. Proteolytic activation of the porcine epidemic diarrhea coronavirus spike fusion protein by trypsin in cell culture. *J Virol* 88:7952–7961. <https://doi.org/10.1128/JVI.00297-14>.
17. Sabir JS, Lam TT, Ahmed MM, Li L, Shen Y, Abo-Aba SE, Qureshi MI, Abu-Zeid M, Zhang Y, Khayami MA, Alharbi NS, Hajrah NH, Sabir MJ, Mutwakil MH, Kabli SA, Alsulaimany FA, Obaid AY, Zhou B, Smith DK, Holmes EC, Zhu H, Guan Y. 2016. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science* 351:81–84. <https://doi.org/10.1126/science.aac8608>.
18. Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, Luo SW, Li PH, Zhang LJ, Guan YJ, Butt KM, Wong KL, Chan KW, Lim W, Shortridge KF, Yuen KY, Peiris JS, Poon LL. 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 302:276–278. <https://doi.org/10.1126/science.1087139>.
19. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwan J, Bhatt S, Peiris JS, Guan Y, Rambaut A. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459:1122–1125. <https://doi.org/10.1038/nature08182>.
20. Ksiazek TG, Rota PA, Rollin PE. 2011. A review of Nipah and Hendra viruses with an historical aside. *Virus Res* 162:173–183. <https://doi.org/10.1016/j.virusres.2011.09.026>.
21. Walls AC, Tortorici MA, Bosch BJ, Frenz B, Rottier PJ, DiMaio F, Rey FA, Veessler D. 2016. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* 531:114–117. <https://doi.org/10.1038/nature16988>.
22. Walls AC, Tortorici MA, Frenz B, Snijder J, Li W, Rey FA, DiMaio F, Bosch BJ, Veessler D. 2016. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* 23:899–905. <https://doi.org/10.1038/nsmb.3293>.
23. Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, Lu G, Wu Y, Yan J, Shi Y, Zhang X, Gao GF. 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* 8:15092. <https://doi.org/10.1038/ncomms15092>.
24. Gui M, Song W, Zhou H, Xu J, Chen S, Xiang Y, Wang X. 2017. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* 27:119–129. <https://doi.org/10.1038/cr.2016.152>.
25. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, Cottrell KS, Graham BS, McLellan JS, Ward AB. 2016. Pre-fusion structure of a human coronavirus spike protein. *Nature* 531:118–121. <https://doi.org/10.1038/nature17200>.
26. Pallesen J, Wang N, Corbett KS, Wrapp D, Kirchdoerfer RN, Turner HL, Cottrell CA, Becker MM, Wang L, Shi W, Kong WP, Andres EL, Kettenbach AN, Denison MR, Chappell JD, Graham BS, Ward AB, McLellan JS. 2017. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc Natl Acad Sci U S A* 114:E7348–E7357. <https://doi.org/10.1073/pnas.1707304114>.
27. Liu C, Tang J, Ma Y, Liang X, Yang Y, Peng G, Qi Q, Jiang S, Li J, Du L, Li F. 2015. Receptor usage and cell entry of porcine epidemic diarrhea coronavirus. *J Virol* 89:6121–6125. <https://doi.org/10.1128/JVI.00430-15>.
28. Hu H, Jung K, Vlasova AN, Chepngeno J, Lu Z, Wang Q, Saif LJ. 2015. Isolation and characterization of porcine deltacoronavirus from pigs with diarrhea in the United States. *J Clin Microbiol* 53:1537–1548. <https://doi.org/10.1128/JCM.00031-15>.
29. Woo PC, Lau SK, Lam CS, Lau CC, Tsang AK, Lau JH, Bai R, Teng JL, Tsang CC, Wang M, Zheng BJ, Chan KH, Yuen KY. 2012. Discovery of seven novel mammalian and avian coronaviruses in the genus *Deltacoronavirus* supports bat coronaviruses as the gene source of *Alphacoronavirus* and *Betacoronavirus* and avian coronaviruses as the gene source of *Gammacoronavirus* and *Deltacoronavirus*. *J Virol* 86:3995–4008. <https://doi.org/10.1128/JVI.06540-11>.
30. Walls A, Tortorici MA, Bosch BJ, Frenz B, Rottier PJ, DiMaio F, Rey FA, Veessler D. 2017. Crucial steps in the structure determination of a coronavirus spike glycoprotein using cryo-electron microscopy. *Protein Sci* 26:113–121. <https://doi.org/10.1002/pro.3048>.
31. Snijder J, Borst AJ, Dosey A, Walls AC, Burrell A, Reddy VS, Kollman JM, Veessler D. 2017. Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids. *J Struct Biol* 198:38–42. <https://doi.org/10.1016/j.jsb.2017.02.008>.
32. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>.
33. DiMaio F, Song Y, Li X, Brunner MJ, Xu C, Conticello V, Egelman E, Marlovits TC, Cheng Y, Baker D. 2015. Atomic-accuracy models from 4.5-Å cryo-electron microscopy data with density-guided iterative local refinement. *Nat Methods* 12:361–365. <https://doi.org/10.1038/nmeth.3286>.
34. Frese CK, Zhou H, Taus T, Altelaar AF, Mechtler K, Heck AJ, Mohammed S. 2013. Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (ETDC). *J Proteome Res* 12:1520–1525. <https://doi.org/10.1021/pr301130k>.
35. Stewart-Jones GB, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, Thomas PV, Wagh K, Zhou T, Behrens AJ, Bylund T, Choi CW, Davison JR, Georgiev IS, Joyce MG, Kwon YD, Pancera M, Taft J, Yang Y, Zhang B, Shivatare SS, Shivatare VS, Lee CD, Wu CY, Bewley CA, Burton DR, Koff WC, Connors M, Crispin M, Baxa U, Korber BT, Wong CH, Mascola JR, Kwong PD. 2016. Trimeric HIV-1-Env structures define glycan shields from clades A, B, and G. *Cell* 165:813–826. <https://doi.org/10.1016/j.cell.2016.04.010>.
36. Gristick HB, von Boehmer L, West AP, Jr, Schamber M, Gazumyan A, Golijanin J, Seaman MS, Fatkenheuer G, Klein F, Nussenzweig MC, Bjorkman PJ. 2016. Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nat Struct Mol Biol* 23:906–915. <https://doi.org/10.1038/nsmb.3291>.
37. Lin YP, Xiong X, Wharton SA, Martin SR, Coombs PJ, Vachieri SG, Christodoulou E, Walker PA, Liu J, Skehel JJ, Gamblin SJ, Hay AJ, Daniels RS, McCauley JW. 2012. Evolution of the receptor binding properties of the

- influenza A(H3N2) hemagglutinin. *Proc Natl Acad Sci U S A* 109: 21474–21479. <https://doi.org/10.1073/pnas.1218841110>.
38. Zhou Y, Lu K, Pfeifferle S, Bertram S, Glowacka I, Drosten C, Pohlmann S, Simmons G. 2010. A single asparagine-linked glycosylation site of the severe acute respiratory syndrome coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple mechanisms. *J Virol* 84:8753–8764. <https://doi.org/10.1128/JVI.00554-10>.
  39. Jeffers SA, Tusell SM, Gillim-Ross L, Hemmila EM, Achenbach JE, Babcock GJ, Thomas WD, Jr, Thackray LB, Young MD, Mason RJ, Ambrosino DM, Wentworth DE, Demartini JC, Holmes KV. 2004. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* 101:15748–15753. <https://doi.org/10.1073/pnas.0403812101>.
  40. Jeffers SA, Hemmila EM, Holmes KV. 2006. Human coronavirus 229E can use CD209L (L-SIGN) to enter cells. *Adv Exp Med Biol* 581:265–269. [https://doi.org/10.1007/978-0-387-33012-9\\_44](https://doi.org/10.1007/978-0-387-33012-9_44).
  41. Krempf C, Ballesteros ML, Zimmer G, Enjuanes L, Klenk HD, Herrler G. 2000. Characterization of the sialic acid binding activity of transmissible gastroenteritis coronavirus by analysis of haemagglutination-deficient mutants. *J Gen Virol* 81:489–496. <https://doi.org/10.1099/0022-1317-81-2-489>.
  42. Sanchez CM, Gebauer F, Sune C, Mendez A, Dopazo J, Enjuanes L. 1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Virology* 190:92–105. [https://doi.org/10.1016/0042-6822\(92\)91195-Z](https://doi.org/10.1016/0042-6822(92)91195-Z).
  43. Huang X, Dong W, Milewska A, Golda A, Qi Y, Zhu QK, Marasco WA, Baric RS, Sims AC, Pyrc K, Li W, Sui J. 2015. Human coronavirus HKU1 spike protein uses  $\alpha$ -acetylated sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme. *J Virol* 89:7202–7213. <https://doi.org/10.1128/JVI.00854-15>.
  44. Vlasak R, Luytjes W, Spaan W, Palese P. 1988. Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc Natl Acad Sci U S A* 85:4526–4529. <https://doi.org/10.1073/pnas.85.12.4526>.
  45. Peng G, Xu L, Lin YL, Chen L, Pasquarella JR, Holmes KV, Li F. 2012. Crystal structure of bovine coronavirus spike protein lectin domain. *J Biol Chem* 287:41931–41938. <https://doi.org/10.1074/jbc.M112.418210>.
  46. Wu K, Li W, Peng G, Li F. 2009. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proc Natl Acad Sci U S A* 106:19970–19974. <https://doi.org/10.1073/pnas.0908837106>.
  47. Reguera J, Santiago C, Mudgal G, Ordone D, Enjuanes L, Casasnovas JM. 2012. Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. *PLoS Pathog* 8:e1002859. <https://doi.org/10.1371/journal.ppat.1002859>.
  48. Ying T, Prabakaran P, Du L, Shi W, Feng Y, Wang Y, Wang L, Li W, Jiang S, Dimitrov DS, Zhou T. 2015. Junctional and allele-specific residues are critical for MERS-CoV neutralization by an exceptionally potent germline-like antibody. *Nat Commun* 6:8223. <https://doi.org/10.1038/ncomms9223>.
  49. Prabakaran P, Gan J, Feng Y, Zhu Z, Choudhry V, Xiao X, Ji X, Dimitrov DS. 2006. Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *J Biol Chem* 281:15829–15836. <https://doi.org/10.1074/jbc.M600697200>.
  50. Hwang WC, Lin Y, Santelli E, Sui J, Jaroszewski L, Stec B, Farzan M, Marasco WA, Liddington RC. 2006. Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. *J Biol Chem* 281:34610–34616. <https://doi.org/10.1074/jbc.M603275200>.
  51. Sui J, Li W, Murakami A, Tamin A, Matthews LJ, Wong SK, Moore MJ, Tallarico AS, Olurinde M, Choe H, Anderson LJ, Bellini WJ, Farzan M, Marasco WA. 2004. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. *Proc Natl Acad Sci U S A* 101:2536–2541. <https://doi.org/10.1073/pnas.0307140101>.
  52. Zhu Z, Chakraborti S, He Y, Roberts A, Sheahan T, Xiao X, Hensley LE, Prabakaran P, Rockx B, Sidorov IA, Corti D, Vogel L, Feng Y, Kim JO, Wang LF, Baric R, Lanzavecchia A, Curtis KM, Nabel GJ, Subbarao K, Jiang S, Dimitrov DS. 2007. Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies. *Proc Natl Acad Sci U S A* 104:12123–12128. <https://doi.org/10.1073/pnas.0701000104>.
  53. Walls AC, Tortorici MA, Snijder J, Xiong X, Bosch B-J, Rey FA, Veerle D. 2017. Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc Natl Acad Sci U S A* 114: 11157–11162. <https://doi.org/10.1073/pnas.1708727114>.
  54. Tse LV, Hamilton AM, Friling T, Whittaker GR. 2014. A novel activation mechanism of avian influenza virus H9N2 by furin. *J Virol* 88:1673–1683. <https://doi.org/10.1128/JVI.02648-13>.
  55. Kawaoka Y, Naeve CW, Webster RG. 1984. Is virulence of H5N2 influenza viruses in chickens associated with loss of carbohydrate from the hemagglutinin? *Virology* 139:303–316. [https://doi.org/10.1016/0042-6822\(84\)90376-3](https://doi.org/10.1016/0042-6822(84)90376-3).
  56. Zheng Q, Deng Y, Liu J, van der Hoek L, Berkhout B, Lu M. 2006. Core structure of S2 from the human coronavirus NL63 spike glycoprotein. *Biochemistry* 45:15205–15215. <https://doi.org/10.1021/bi061686w>.
  57. Matsuyama S, Ujike M, Morikawa S, Tashiro M, Taguchi F. 2005. Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection. *Proc Natl Acad Sci U S A* 102:12543–12547. <https://doi.org/10.1073/pnas.0503203102>.
  58. Matsuyama S, Taguchi F. 2009. Two-step conformational changes in a coronavirus envelope glycoprotein mediated by receptor binding and proteolysis. *J Virol* 83:11133–11141. <https://doi.org/10.1128/JVI.00959-09>.
  59. Eckert DM, Malashkevich VN, Kim PS. 1998. Crystal structure of GCN4-pIQ1, a trimeric coiled coil with buried polar residues. *J Mol Biol* 284: 859–865. <https://doi.org/10.1006/jmbi.1998.2214>.
  60. Yin HS, Wen X, Paterson RG, Lamb RA, Jardetzky TS. 2006. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* 439:38–44. <https://doi.org/10.1038/nature04322>.
  61. Berrow NS, Alderton D, Sainsbury S, Nettleship J, Assenberg R, Rahman N, Stuart DI, Owens RJ. 2007. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res* 35:e45. <https://doi.org/10.1093/nar/gkm047>.
  62. Mastrorade DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* 152: 36–51. <https://doi.org/10.1016/j.jsb.2005.07.007>.
  63. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14:331–332. <https://doi.org/10.1038/nmeth.4193>.
  64. Zhang K. 2016. Gctf: real-time CTF determination and correction. *J Struct Biol* 193:1–12. <https://doi.org/10.1016/j.jsb.2015.11.003>.
  65. Rohou A, Grigorieff N. 2015. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192:216–221. <https://doi.org/10.1016/j.jsb.2015.08.008>.
  66. Voss NR, Yoshioka CK, Radermacher M, Potter CS, Carragher B. 2009. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J Struct Biol* 166:205–213. <https://doi.org/10.1016/j.jsb.2009.01.004>.
  67. Kimanius D, Forsberg BO, Scheres SH, Lindahl E. 2016. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* 5:e18722.
  68. Scheres SH. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180:519–530. <https://doi.org/10.1016/j.jsb.2012.09.006>.
  69. Scheres SH, Chen S. 2012. Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9:853–854. <https://doi.org/10.1038/nmeth.2115>.
  70. Rosenthal PB, Henderson R. 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* 333:721–745. <https://doi.org/10.1016/j.jmb.2003.07.013>.
  71. Chen S, McMullan G, Faruqi AR, Murshudov GN, Short JM, Scheres SH, Henderson R. 2013. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 135:24–35. <https://doi.org/10.1016/j.ultramic.2013.06.004>.
  72. Goddard TD, Huang CC, Ferrin TE. 2007. Visualizing density maps with UCSF Chimera. *J Struct Biol* 157:281–287. <https://doi.org/10.1016/j.jsb.2006.06.010>.
  73. Brown A, Long F, Nicholls RA, Toots J, Emsley P, Murshudov G. 2015. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr D Biol Crystallogr* 71: 136–153. <https://doi.org/10.1107/S1399004714021683>.
  74. DiMaio F, Leaver-Fay A, Bradley P, Baker D, Andre I. 2011. Modeling symmetric macromolecular structures in Rosetta3. *PLoS One* 6:e20450. <https://doi.org/10.1371/journal.pone.0020450>.

75. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12–21. <https://doi.org/10.1107/S0907444909042073>.
76. Agirre J, Iglesias-Fernandez J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. 2015. Privateer: software for the conformational validation of carbohydrate structures. *Nat Struct Mol Biol* 22:833–834. <https://doi.org/10.1038/nsmb.3115>.
77. Bern M, Kil YJ, Becker C. 2012. Byonic: advanced peptide and protein identification software. *Curr Protoc Bioinformatics* Chapter 13:Unit 13.20.